

OIPE

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/841,720A

DATE: 12/05/2001

TIME: 15:07:25

Input Set : N:\Crf3\RULE60\09841720A.RAW

Output Set: N:\CRF3\12052001\I841720A.raw

SEQUENCE LISTING

3 (1) GENERAL INFORMATION:

5 (i) APPLICANT: Yu, Lei

7 (ii) TITLE OF INVENTION: MU OPIOID RECEPTORS: COMPOSOTIONS AND
8 METHODS

10 (iii) NUMBER OF SEQUENCES: 9

12 (iv) CORRESPONDENCE ADDRESS:

13 (A) ADDRESSEE: Arnold, White & Durkee

14 (B) STREET: P.O. Box 4433

15 (C) CITY: Houston

16 (D) STATE: Texas

17 (E) COUNTRY: USA

18 (F) ZIP: 77210

20 (v) COMPUTER READABLE FORM:

21 (A) MEDIUM TYPE: Floppy disk

22 (B) COMPUTER: IBM PC compatible

23 (C) OPERATING SYSTEM: PC-DOS/MS-DOS

24 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30

26 (vi) CURRENT APPLICATION DATA:

C--> 27 (A) APPLICATION NUMBER: US/09/841,720A

C--> 28 (B) FILING DATE: 24-Apr-2001

29 (C) CLASSIFICATION: 435

31 (vii) PRIOR APPLICATION DATA:

32 (A) APPLICATION NUMBER: 08/120,601

33 (B) FILING DATE: 1993-09-13

35 (viii) ATTORNEY/AGENT INFORMATION:

36 (A) NAME: Wilson, Mark B.

37 (B) REGISTRATION NUMBER: 37,259

38 (C) REFERENCE/DOCKET NUMBER: INDA:002

40 (ix) TELECOMMUNICATION INFORMATION:

41 (A) TELEPHONE: 512/418-3000

42 (B) TELEFAX: 512/474-7577

45 (2) INFORMATION FOR SEQ ID NO: 1:

47 (i) SEQUENCE CHARACTERISTICS:

48 (A) LENGTH: 1618 base pairs

49 (B) TYPE: nucleic acid

50 (C) STRANDEDNESS: single

51 (D) TOPOLOGY: linear

54 (ix) FEATURE:

55 (A) NAME/KEY: CDS

56 (B) LOCATION: 214..1407

59 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

61 CGTGGAAGGG GGCTACAAGC AGAGGAGAAT ATCAGACGCT CAGACGTTCC CTTCTGCCTG 60

63 CCGCTCTTCT CTGGTTCCAC TAGGGCTGGT CCATGTAAGA ATCTGACGGA GCCTAGGGCA 120

65 GCTGTGAGAG GAAGAGGCTG GGGCGCGTGG AACCCGAAAA GTCTGAGTGC TCTCAGTTAC 180

67 AGCCTACCTA GTCCGCAGCA GGCCTTCAGC ACC ATG GAC AGC AGC ACC GGC CCA 234

68 Met Asp Ser Ser Thr Gly Pro

ENTERED

RAW SEQUENCE LISTING

DATE: 12/05/2001

PATENT APPLICATION: US/09/841,720A

TIME: 15:07:25

Input Set : N:\Crf3\RULE60\09841720A.RAW

Output Set: N:\CRF3\12052001\I841720A.raw

69										1				5				
71	GGG	AAC	ACC	AGC	GAC	TGC	TCA	GAC	CCC	TTA	GCT	CAG	GCA	AGT	TGC	TCC		282
72	Gly	Asn	Thr	Ser	Asp	Cys	Ser	Asp	Pro	Leu	Ala	Gln	Ala	Ser	Cys	Ser		
73			10						15					20				
75	CCA	GCA	CCT	GGC	TCC	TGG	CTC	AAC	TTG	TCC	CAC	GTT	GAT	GGC	AAC	CAG		330
76	Pro	Ala	Pro	Gly	Ser	Trp	Leu	Asn	Leu	Ser	His	Val	Asp	Gly	Asn	Gln		
77		25					30					35						
79	TCC	GAT	CCA	TGC	GGT	CTG	AAC	CGC	ACC	GGG	CTT	GGC	GGG	AAC	GAC	AGC		378
80	Ser	Asp	Pro	Cys	Gly	Leu	Asn	Arg	Thr	Gly	Leu	Gly	Gly	Asn	Asp	Ser		
81	40					45				50				55				
83	CTG	TGC	CCT	CAG	ACC	GGC	AGC	CCT	TCC	ATG	GTC	ACA	GCC	ATT	ACC	ATC		426
84	Leu	Cys	Pro	Gln	Thr	Gly	Ser	Pro	Ser	Met	Val	Thr	Ala	Ile	Thr	Ile		
85				60					65					70				
87	ATG	GCC	CTC	TAC	TCT	ATC	GTG	TGT	GTA	GTG	GGC	CTC	TTC	GGA	AAC	TTC		474
88	Met	Ala	Leu	Tyr	Ser	Ile	Val	Cys	Val	Val	Gly	Leu	Phe	Gly	Asn	Phe		
89			75					80				85						
91	CTG	GTC	ATG	TAT	GTG	ATT	GTA	AGA	TAC	ACC	AAA	ATG	AAG	ACT	GCC	ACC		522
92	Leu	Val	Met	Tyr	Val	Ile	Val	Arg	Tyr	Thr	Lys	Met	Lys	Thr	Ala	Thr		
93			90					95				100						
95	AAC	ATC	TAC	ATT	TTC	AAC	CTT	GCT	CTG	GCA	GAC	GCC	TTA	GCG	ACC	AGT		570
96	Asn	Ile	Tyr	Ile	Phe	Asn	Leu	Ala	Leu	Ala	Asp	Ala	Leu	Ala	Thr	Ser		
97		105				110				115								
99	ACA	CTG	CCC	TTT	CAG	AGT	GTC	AAC	TAC	CTG	ATG	GGA	ACA	TGG	CCC	TTC		618
100	Thr	Leu	Pro	Phe	Gln	Ser	Val	Asn	Tyr	Leu	Met	Gly	Thr	Trp	Pro	Phe		
101	120					125				130				135				
103	GGA	ACC	ATC	CTC	TGC	AAG	ATC	GTG	ATC	TCA	ATA	GAT	TAC	TAC	AAC	ATG		666
104	Gly	Thr	Ile	Leu	Cys	Lys	Ile	Val	Ile	Ser	Ile	Asp	Tyr	Tyr	Asn	Met		
105				140					145				150					
107	TTC	ACC	AGC	ATA	TTC	ACC	CTC	TGC	ACC	ATG	AGC	GTG	GAC	CGC	TAC	ATT		714
108	Phe	Thr	Ser	Ile	Phe	Thr	Leu	Cys	Thr	Met	Ser	Val	Asp	Arg	Tyr	Ile		
109			155					160				165						
111	GCT	GTC	TGC	CAC	CCA	GTC	AAA	GCC	CTG	GAT	TTC	CGT	ACC	CCC	CGA	AAT		762
112	Ala	Val	Cys	His	Pro	Val	Lys	Ala	Leu	Asp	Phe	Arg	Thr	Pro	Arg	Asn		
113		170					175				180							
115	GCC	AAA	ATC	GTC	AAC	GTC	TGC	AAC	TGG	ATC	CTC	TCT	TCT	GCC	ATC	GGT		810
116	Ala	Lys	Ile	Val	Asn	Val	Cys	Asn	Trp	Ile	Leu	Ser	Ser	Ala	Ile	Gly		
117		185				190				195								
119	CTG	CCT	GTA	ATG	TTC	ATG	GCA	ACC	ACA	AAA	TAC	AGG	CAG	GGG	TCC	ATA		858
120	Leu	Pro	Val	Met	Phe	Met	Ala	Thr	Thr	Lys	Tyr	Arg	Gln	Gly	Ser	Ile		
121	200				205					210				215				
123	GAT	TGC	ACC	CTC	ACG	TTC	TCC	CAC	CCA	ACC	TGG	TAC	TGG	GAG	AAC	CTG		906
124	Asp	Cys	Thr	Leu	Thr	Phe	Ser	His	Pro	Thr	Trp	Tyr	Trp	Glu	Asn	Leu		
125				220						225				230				
127	CTC	AAA	ATC	TGT	GTC	TTT	ATC	TTC	GCT	TTC	ATC	ATG	CCG	ATC	CTC	ATC		954
128	Leu	Lys	Ile	Cys	Val	Phe	Ile	Phe	Ala	Phe	Ile	Met	Pro	Ile	Leu	Ile		
129			235					240				245						
131	ATC	ACT	GTG	TGT	TAC	GGC	CTG	ATG	ATC	TTA	CGA	CTC	AAG	AGC	GTT	CGC		1002
132	Ile	Thr	Val	Cys	Tyr	Gly	Leu	Met	Ile	Leu	Arg	Leu	Lys	Ser	Val	Arg		
133			250				255					260						

RAW SEQUENCE LISTING

DATE: 12/05/2001

PATENT APPLICATION: US/09/841,720A

TIME: 15:07:25

Input Set : N:\Crf3\RULE60\09841720A.RAW

Output Set: N:\CRF3\12052001\I841720A.raw

```

135 ATG CTA TCG GGC TCC AAA GAA AAG GAC AGG AAT CTG CGC AGG ATC ACC      1050
136 Met Leu Ser Gly Ser Lys Glu Lys Asp Arg Asn Leu Arg Arg Ile Thr
137      265                      270                      275
139 CGG ATG GTG CTG GTG GTC GTG GCT GTA TTT ATC GTC TGC TGG ACC CCC      1098
140 Arg Met Val Leu Val Val Val Ala Val Phe Ile Val Cys Trp Thr Pro
141 280                      285                      290                      295
143 ATC CAC ATC TAC GTC ATC ATC AAA GCG CTG ATC ACG ATT CCA GAA ACC      1146
144 Ile His Ile Tyr Val Ile Ile Lys Ala Leu Ile Thr Ile Pro Glu Thr
145                      300                      305                      310
147 ACA TTT CAG ACC GTT TCC TGG CAC TTC TGC ATT GCT TTG GGT TAC ACG      1194
148 Thr Phe Gln Thr Val Ser Trp His Phe Cys Ile Ala Leu Gly Tyr Thr
149                      315                      320                      325
151 AAC AGC TGC CTG AAT CCA GTT CTT TAC GCC TTC CTG GAT GAA AAC TTC      1242
152 Asn Ser Cys Leu Asn Pro Val Leu Tyr Ala Phe Leu Asp Glu Asn Phe
153      330                      335                      340
155 AAG CGA TGC TTC AGA GAG TTC TGC ATC CCA ACC TCG TCC ACG ATC GAA      1290
156 Lys Arg Cys Phe Arg Glu Phe Cys Ile Pro Thr Ser Ser Thr Ile Glu
157      345                      350                      355
159 CAG CAA AAC TCC ACT CGA GTC CGT CAG AAC ACT AGG GAA CAT CCC TCC      1338
160 Gln Gln Asn Ser Thr Arg Val Arg Gln Asn Thr Arg Glu His Pro Ser
161 360                      365                      370                      375
163 ACG GCT AAT ACA GTG GAT CGA ACT AAC CAC CAG CTA GAA AAT CTG GAG      1386
164 Thr Ala Asn Thr Val Asp Arg Thr Asn His Gln Leu Glu Asn Leu Glu
165      380                      385                      390
167 GCA GAA ACT GCT CCA TTG CCC TAACTGGGTC TCACACCATC CAGACCCTCG      1437
168 Ala Glu Thr Ala Pro Leu Pro
169      395
171 CTAAGCTTAG AGGCCGCCAT CTACGTGGAA TCAGGTTGCT GTCAGGGTGT GTGGGAGGCT      1497
173 CTGGTTTCCT GAGAAACCAT CTGATCCTGC ATTCAAAAGTC ATTCCTCTCT GGCTACTTCA      1557
175 CTCTGCACAT GAGAGATGCT CAGACTGATC AAGACCAGAA GAAAGAAGAG ACTACCGGAC      1617
177 A      1618
180 (2) INFORMATION FOR SEQ ID NO: 2:
182      (i) SEQUENCE CHARACTERISTICS:
183          (A) LENGTH: 398 amino acids
184          (B) TYPE: amino acid
185          (D) TOPOLOGY: linear
187      (ii) MOLECULE TYPE: protein
189      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
191 Met Asp Ser Ser Thr Gly Pro Gly Asn Thr Ser Asp Cys Ser Asp Pro
192  1                      5                      10                      15
194 Leu Ala Gln Ala Ser Cys Ser Pro Ala Pro Gly Ser Trp Leu Asn Leu
195      20                      25                      30
197 Ser His Val Asp Gly Asn Gln Ser Asp Pro Cys Gly Leu Asn Arg Thr
198      35                      40                      45
200 Gly Leu Gly Gly Asn Asp Ser Leu Cys Pro Gln Thr Gly Ser Pro Ser
201      50                      55                      60
203 Met Val Thr Ala Ile Thr Ile Met Ala Leu Tyr Ser Ile Val Cys Val
204  65                      70                      75                      80
206 Val Gly Leu Phe Gly Asn Phe Leu Val Met Tyr Val Ile Val Arg Tyr

```

RAW SEQUENCE LISTING

DATE: 12/05/2001

PATENT APPLICATION: US/09/841,720A

TIME: 15:07:25

Input Set : N:\Crf3\RULE60\09841720A.RAW

Output Set: N:\CRF3\12052001\I841720A.raw

```

207          85          90          95
209 Thr Lys Met Lys Thr Ala Thr Asn Ile Tyr Ile Phe Asn Leu Ala Leu
210          100          105          110
212 Ala Asp Ala Leu Ala Thr Ser Thr Leu Pro Phe Gln Ser Val Asn Tyr
213          115          120          125
215 Leu Met Gly Thr Trp Pro Phe Gly Thr Ile Leu Cys Lys Ile Val Ile
216          130          135          140
218 Ser Ile Asp Tyr Tyr Asn Met Phe Thr Ser Ile Phe Thr Leu Cys Thr
219 145          150          155          160
221 Met Ser Val Asp Arg Tyr Ile Ala Val Cys His Pro Val Lys Ala Leu
222          165          170          175
224 Asp Phe Arg Thr Pro Arg Asn Ala Lys Ile Val Asn Val Cys Asn Trp
225          180          185          190
227 Ile Leu Ser Ser Ala Ile Gly Leu Pro Val Met Phe Met Ala Thr Thr
228          195          200          205
230 Lys Tyr Arg Gln Gly Ser Ile Asp Cys Thr Leu Thr Phe Ser His Pro
231          210          215          220
233 Thr Trp Tyr Trp Glu Asn Leu Leu Lys Ile Cys Val Phe Ile Phe Ala
234 225          230          235          240
236 Phe Ile Met Pro Ile Leu Ile Ile Thr Val Cys Tyr Gly Leu Met Ile
237          245          250          255
239 Leu Arg Leu Lys Ser Val Arg Met Leu Ser Gly Ser Lys Glu Lys Asp
240          260          265          270
242 Arg Asn Leu Arg Arg Ile Thr Arg Met Val Leu Val Val Ala Val
243          275          280          285
245 Phe Ile Val Cys Trp Thr Pro Ile His Ile Tyr Val Ile Ile Lys Ala
246          290          295          300
248 Leu Ile Thr Ile Pro Glu Thr Thr Phe Gln Thr Val Ser Trp His Phe
249 305          310          315          320
251 Cys Ile Ala Leu Gly Tyr Thr Asn Ser Cys Leu Asn Pro Val Leu Tyr
252          325          330          335
254 Ala Phe Leu Asp Glu Asn Phe Lys Arg Cys Phe Arg Glu Phe Cys Ile
255          340          345          350
257 Pro Thr Ser Ser Thr Ile Glu Gln Gln Asn Ser Thr Arg Val Arg Gln
258          355          360          365
260 Asn Thr Arg Glu His Pro Ser Thr Ala Asn Thr Val Asp Arg Thr Asn
261          370          375          380
263 His Gln Leu Glu Asn Leu Glu Ala Glu Thr Ala Pro Leu Pro
264 385          390          395

```

267 (2) INFORMATION FOR SEQ ID NO: 3:

269 (i) SEQUENCE CHARACTERISTICS:

270 (A) LENGTH: 1618 base pairs

271 (B) TYPE: nucleic acid

272 (C) STRANDEDNESS: single

273 (D) TOPOLOGY: linear

276 (ix) FEATURE:

277 (A) NAME/KEY: CDS

278 (B) LOCATION: 339..1232

281 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

RAW SEQUENCE LISTING

DATE: 12/05/2001

PATENT APPLICATION: US/09/841,720A

TIME: 15:07:25

Input Set : N:\Crf3\RULE60\09841720A.RAW

Output Set: N:\CRF3\12052001\I841720A.raw

```

283 CGTGAAGGG GGCTACAAGC AGAGGAGAAT ATCAGACGCT CAGACGTTCC CTTCTGCCTG      60
285 CCGCTCTTCT CTGGTTCCAC TAGGGCTGGT CCATGTAAGA ATCTGACGGA GCCTAGGGCA      120
287 GCTGTGAGAG GAAGAGGCTG GGGCGCGTGG AACCCGAAAA GTCTGAGTGC TCTCAGTTAC      180
289 AGCCTACCTA GTCCGACAGC GGCCTTCAGC ACCATGGACA GCAGCACCGG CCCAGGGAAC      240
291 ACCAGCGACT GCTCAGACCC CTTAGCTCAG GCAAGTTGCT CCCAGCACCC TGGCTCCTGG      300
293 CTCAACTTGT CCCACGTTGA TGGCAACCAG TCCGATCC ATG CGG TCT GAA CCG      353
294                                     Met Arg Ser Glu Pro
295                                     1           5
297 CAC CGG GCT TGG CGG GAA CGA CAG CCT GTG CCC TCA GAC CGG CAG CCC      401
298 His Arg Ala Trp Arg Glu Arg Gln Pro Val Pro Ser Asp Arg Gln Pro
299             10             15             20
301 TTC CAT GGT CAC AGC CAT TAC CAT CAT GGC CCT CTA CTC TAT CGT GTG      449
302 Phe His Gly His Ser His Tyr His His Gly Pro Leu Leu Tyr Arg Val
303             25             30             35
305 TGT AGT GGG CCT CTT CGG AAA CTT CCT GGT CAT GTA TGT GAT TGT AAG      497
306 Cys Ser Gly Pro Leu Arg Lys Leu Pro Gly His Val Cys Asp Cys Lys
307             40             45             50
309 ATA CAC CAA AAT GAA GAC TGC CAC CAA CAT CTA CAT TTT CAA CCT TGC      545
310 Ile His Gln Asn Glu Asp Cys His Gln His Leu His Phe Gln Pro Cys
311             55             60             65
313 TCT GGC AGA CGC CTT AGC GAC CAG TAC ACT GCC CTT TCA GAG TGT CAA      593
314 Ser Gly Arg Arg Leu Ser Asp Gln Tyr Thr Ala Leu Ser Glu Cys Gln
315             70             75             80             85
317 CTA CCT GAT GGG AAC ATG GCC CTT CGG AAC CAT CCT CTG CAA GAT CGT      641
318 Leu Pro Asp Gly Asn Met Ala Leu Arg Asn His Pro Leu Gln Asp Arg
319             90             95             100
321 GAT CTC AAT AGA TTA CTA CAA CAT GTT CAC CAG CAT ATT CAC CCT CTG      689
322 Asp Leu Asn Arg Leu Leu Gln His Val His Gln His Ile His Pro Leu
323             105             110             115
325 CAC CAT GAG CGT GGA CCG CTA CAT TGC TGT CTG CCA CCC AGT CAA AGC      737
326 His His Glu Arg Gly Pro Leu His Cys Cys Leu Pro Pro Ser Gln Ser
327             120             125             130
329 CCT GGA TTT CCG TAC CCC CCG AAA TGC CAA AAT CGT CAA CGT CTG CAA      785
330 Pro Gly Phe Pro Tyr Pro Pro Lys Cys Gln Asn Arg Gln Arg Leu Gln
331             135             140             145
333 CTG GAT CCT CTC TTC TGC CAT CGG TCT GCC TGT AAT GTT CAT GGC AAC      833
334 Leu Asp Pro Leu Phe Cys His Arg Ser Ala Cys Asn Val His Gly Asn
335             150             155             160             165
337 CAC AAA ATA CAG GCA GGG GTC CAT AGA TTG CAC CCT CAC GTT CTC CCA      881
338 His Lys Ile Gln Ala Gly Val His Arg Leu His Pro His Val Leu Pro
339             170             175             180
341 CCC AAC CTG GTA CTG GGA GAA CCT GCT CAA AAT CTG TGT CTT TAT CTT      929
342 Pro Asn Leu Val Leu Gly Glu Pro Ala Gln Asn Leu Cys Leu Tyr Leu
343             185             190             195
345 CGC TTT CAT CAT GCC GAT CCT CAT CAT CAC TGT GTG TTA CGG CCT GAT      977
346 Arg Phe His His Ala Asp Pro His His His Cys Val Leu Arg Pro Asp
347             200             205             210
349 GAT CTT ACG ACT CAA GAG CGT TCG CAT GCT ATC GGG CTC CAA AGA AAA      1025
350 Asp Leu Thr Thr Gln Glu Arg Ser His Ala Ile Gly Leu Gln Arg Lys

```

VERIFICATION SUMMARY

DATE: 12/05/2001

PATENT APPLICATION: US/09/841,720A

TIME: 15:07:26

Input Set : N:\Crf3\RULE60\09841720A.RAW

Output Set: N:\CRF3\12052001\I841720A.raw

L:27 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]

L:28 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]